Explicit Routing with QoS Constraints in IP over WDM

SuKyoung Lee, David Griffith, Vincent Coussot and David Su

National Institute of Standards and Technology

Advanced Network Technologies Division, Room 467

100 Bureau Drive, Stop 8920, Gaithersburg, MD 20899, USA

Email: sukyoung@antd.nist.gov

Abstract

Generalized Multiprotocol Label Switching (GMPLS) is being developed as a unified control architecture for the optical transport network and the layers above it. GMPLS supports Traffic Engineering (TE) by allowing Explicit Routing (ER) of data-bearing paths across networks, which will help to guarantee Quality of Service (QoS) for types of services in IP over WDM networks. Although there have been many approaches to allocate resources in optical data networks using constrained linear programs, these do not consider the delays at layer 3, which impacts QoS. In networks where it is not possible to create a virtual topology at layer 3 that is a full mesh, more than one lightpath will be required to route traffic between certain pairs of source and destination IP routers. In this paper, we present a technique for traffic engineering in optical networks that support QoS considering the traffic flows with delay QoS requirements across optical networks. This technique provides real-time services with a specific optical label switched path and optimizes a objective function including the queueing delay at layer 3.

Keywords: Optical Networks, Traffic Engineering, Explicit Routing, Multi-Service QoS

1 Introduction

As the global telecommunications infrastructure has expanded, the traditional distinction between circuit-switched networks and packet-switched networks has become blurred. This phenomenon is partially due to the development of mechanisms to facilitate the transport of real-time, delay-sensitive data over packet-switched networks. It is also due to the strong growth in data traffic, which recently caused data to surpass voice as the dominant form of information carried on the world's networks. This explosive increase rate of Internet traffic volumes poses a challenging scalability problem for traffic carriers. Optical network technologies using Wavelength Division Multiplexing (WDM) offer one potent solution to this problem. Moreover, research is ongoing to introduce more intelligence in the control plane of optical transport systems which will make them more survivable, flexible, controllable and open for traffic engineering. Recently, there has been considerable work on the efficient internetworking of higher layers, primarily the IP layer, with the WDM-capable optical layer. In this lambda labeling network, a given wavelength channel on an incoming link to an optical switch can be routed across an optical switching fabric to another wavelength channel on an outgoing link. The overall operation is of a circuit-

switched nature because the wavelength switching associations are determined at setup time. The provisioning and restoration of end-to-end optical trails would be supported in this network.

In MPLS [1, 2], arriving packets are mapped to FECs (Forwarding Equivalence Classes) at the ingress routers of an MPLS network. Based on these FECs, each switch provisions resources for a LSP (Label Switched Path) by establishing mappings from labels received from downstream (toward destination) LSRs (Label Switching Routers) to output ports. This idea can be extended to optical networks, in which wavelengths have the same role that labels do in an MPLS network. In the optical case, a GMPLS label mapping is used by OXCs (Optical Cross-connects) to establish a switching connection between an (input port, input lambda) tuple and an (output port, output lambda) tuple in the cross-connect table. An important reason for using GMPLS is that it can be used as a powerful tool for TE (Traffic Engineering). The goal of TE is to make best use of the network infrastructure and this is facilitated by the ER feature of MPLS. TE can also be used to support QoS (Quality of Service) by distributing different traffic types in the network such that their respective delay, delay variation, and loss requirements are met. There have been several efforts directed towards using QoS information to make routing decisions. A fairly straightforward approach is to use delay and bandwidth measures to compute link costs, as defined in [3]. The link costs can be propagated throughout the network using mechanisms such as the OSPF (Open Shortest Path First) flooding protocol.

Routing that accounts for QoS constraints can be extended to optical networks, as discussed in [4]. This approach requires a means of routing the lightpaths and LSPs within them so that the QoS constraints of the carried traffic are met. The usual offline routing mechanisms that make links between certain ingress and egress pairs can get congested although other (possibly underutilized) links are available on alternate paths. This results in an unnecessary higher delay for some traffic while resources elsewhere in the network go unused. In this paper, we propose a mechanism to provide better delay QoS in the optical network by efficiently utilizing the available wavelengths. We do this by using a linear programming approach that seeks to minimize the total path delay, including delay introduced by packet processing at layers above optical. We used the flow constraints discussed in [5]. While other researchers, specifically [6], have introduced linear programs that account for the bottleneck effect at layer 3, their program seeks to maximize total network throughput, while ours is focused on meeting QoS requirements.

Optical channel trails can be specified via ER, which allows ISPs (Internet Service Providers) or large carriers to engineer the traffic flows in their own networks. The traditional Internet that supported only best effort service is being transformed very quickly into a commercial environment where users (and, therefore, ISPs and large carriers) are increasingly demanding support for various types of QoS (Quality of Service). Furthermore new compound services such as video-conferencing and especially Internet telephony have to be provided with the required QoS even as network utilization increases. In order to support multiple service types in an IP over WDM network, we propose a TE mechanism using MPLS that incorporates a delay-based QoS metric. We will formulate the problem as an optimization problem, using principles from multi-commodity flow for physical routing of lightpaths and traffic flows on the virtual topology.

The optimization routine that we are proposing accounts for path delay, including queueing delay at layer 3. In an optical network using lambda labeling, processing in layer 3 would constitute a significant bottleneck whose impact we wish to minimize. In optical networks in which it is possible to create lightpath connections between every pair of edge routers, this is not an issue, but in large networks with thousands of devices at the edge, creating

2

a virtual (lightpath) topology being a full mesh is impractical. In such cases, we are more likely to have a virtual topology like that depicted in Fig. 1. In this example network, a direct connection can be established between Router A and Router E, for instance, but a connection between Router A and Router C would require the use of two lightpaths joined at either Router B or Router D. Additionally, the future optical network architecture as described in [7] allows for a non-homogeneous optical core that may be subdivided into multiple administrative regions. If we are supporting DiffServ (Differentiated Services) over such a core network, it may be necessary to recompute DiffServ codepoints at the administrative boundaries, since different network operators may use different codepoint sets, associate different Per-Hop Behaviors with a given codepoint, or not support DiffServ at all. Because of these factors, a given path may pass through intermediate nodes with layer 3 functionality in a lambda labeling network; it is thus useful to consider the information of all resources with respect to layer 3, in addition to accounting for those at layer 2, when the route server decides the explicit route. The layer 3 processing overhead from intermediate node is reflected into the proposed path set-up algorithm. Moreover, while other researchers, specifically [6], have introduced linear program that accounts for the bottleneck effect at layer3 and seeks to maximize total network throughput, ours is focused on meeting QoS requirements.

In future optical networks that are managed using a combination of GMPLS and higher-layer IP-based protocols, certain OXCs with interfaces located between optical subnetworks, will employ some kind of layer 3 processing for some of the wavelengths in the data stream. This type of traffic grooming means that the user traffic will be subjected to queueing delays that can affect service quality. For this reason, LSP computation should account for effects at higher layers that can impact QoS in addition to optical layer parameters such as bit error rate and line delay.

This paper is organized as follows. We first provide a general discussion of routing issues in optical networks and the additional constraints that are imposed by offering support for various grades of QoS. Next we give a formulation of the ER-LSP routing problem as a linear program; we specify the objective function to be minimized and list the relevant constraints. The program will then choose a path over a given virtual topology that meets the delay criterion in the new flow's traffic specification. We then demonstrate the utility of this approach with some simulation runs that we performed using a sample logical topology.

2   Routing and QoS Provisioning in Lambda Labeling Network

The optical core network consists of one or more optical subnetworks, each of which is administered by a single entity such as a large carrier. The subnetworks are, in general, composed of a mixture of transparent optical switches (known as Photonic Cross-connects (PXCs)) that do not perform any Optical/Electrical/Optical (O/E/O) conversions and opaque Optical Cross-Connects (OXCs) that carry out some type of O/E/O operations. Some OXCs may also incorporate higher-layer processing functions, particularly support for the IP and MPLS protocol suites, as well as Asynchronous Transfer Mode (ATM) and the Synchronous Digital Hierarchy/Synchronous Optical NETwork (SDH/SONET) [8, 9]. We refer to these OXCs as Optical-Lambda Switching Routers (O-LSRs). They can be viewed as a combination of a router and an OXC. The routing function can be built-in to the OXC or it can reside in a separate piece of equipment. The IP router is responsible for all the layer 3 functions such as addressing, routing, and global topology discovery. It is also responsible for optimizing network performance, which can be carried out via TE with QoS support, management of optical resources (i.e. wavelength assignment in coordination with the optical channel sublayer), and restoration. Each OXC is capable of switching a data

stream from a given input port to a given output port by appropriately configuring an internal crossconnect table. A lightpath is established by setting up suitable crossconnects in the ingress, egress and a set of intermediate OXCs such as that a continuous physical path exists across the optical network.

The passive or configurable OXCs and their fiber connections constitute the physical topology of the network. The virtual (or logical) topology consists of the lightpaths across the network and is determined by the configuration of the optical Add/Drop Multiplexers (ADMs) and transmitters and receivers on each node. That is, the virtual topology is the topology seen by the higher layers that use the optical layer.

We propose a mechanism to support multiple services transport in lambda labeling capable networks. The virtual topology (virtual network) is made by routing the lightpaths over the physical topology. Then, the wavelengths are assigned dynamically to the lightpaths for multiple service classes. When a new flow is to be routed through the network, an ingress O-LSR determines the virtual path it will be routed through, in terms of the QoS requirements of the flow, such as the maximum acceptable delay.

If a set of lightpath requests is known, the network's resources can be utilized most efficiently by employing an offline routing algorithm, in which all the lightpaths are routed and assigned wavelengths at the same time, in a manner such that the resources consumed are minimized. In the emerging business models that support bandwidth on demand services, for instance, it may not be possible to do offline batch processing of lightpath and LSP setup requests. Also, the QoS requirements of existing requests may change over time. In such a situation, an online TE algorithm that can accommodate the QoS requirements of new requests is needed.

We consider priority as another important element for TE. Although, in this paper, two levels of priority are taken into account where priority access to wavelength is given to the QoS service class over BE (Best Effort) traffic, the proposed mechanism extends to multiple priority levels. In an lambda labeling capable IP network of considerable size, two different service classes are assumed to be supported: DS (Delay Sensitive), and BE service classes [10]. (In [10], TS (Throughput Sensitive) traffic is also included, but we consider only the above two services in this paper.)

In the following paragraphs, we describe the major elements of the Route Server architecture that are used to support the provisioning operation.

Packet Classifier
The classification of packets into FECs is done using a packet classifier that examines header fields such as source address, destination address, ToS (Type of Service) (if DiffServ is supported), and others. In other words, the packets are first classified at the ingress O-LSR. Then a mapping between the FEC and a LSP must take place. This is done by providing a FEC specification for each LSP. The FECs could be determined in different ways such as by interaction with routing protocols or by a TE server. These classification rules can be downloaded from a policer in a TE server for the network. This classification enables the network operator to engineer the traffic in the network and route each FEC in a specified manner.

TE server
ER allows routes to be specified using centralized, distributed, or even hybrid computational frameworks. In the case of a centralized server, the policer should also be centralized and the server must be able to consolidate the topology and link state information for the entire network. It then uses this information to compute ERs in

response to requests from the ingress O-LSRs. On the other hand, when TE is done in a distributed manner, an ingress O-LSR computes an explicit route without communicating with a TE server. Each ingress node must then support a local TE procedure. When a hybrid approach is used, the centralized TE server and the local TE processes will interact closely so as to coordinate the wavelength assignment on the virtual network.

The centralized TE server periodically requests OXCs to send their traffic conditions. How often this occurs (hourly or daily, for example) depends on the network's administrative policies. After receiving request message from the server, each OXC sends its traffic state information. Then, the TE server analyzes the received information and sets up new ER-LSPs or reconfigures some lightpaths in the current network topology according to that information so as to optimize the overall network performance.

Signaling

When ER is applied to lambda labeling, the LSPs can be set up by specifying the IP addresses of the MPLS nodes along the route in the ER TLV (Type-Length-Value) object, which is supported by both the RSVP-TE (Resource ReServation Protocol with TE Extensions) and CR (Constrained Routing)-LDP label distribution protocols [13],[14]. These LSPs can actually be lightpath channels between O-LSR nodes. A connection request from a source is received by the first hop router (an ingress O-LSR). This router creates a lightpath setup request message and sends it towards the destination of the lightpath. The message is received by the last hop node on the default routed lightpath as the payload of a normal IP packet. Then, a wavelength is assigned for the lightpath at every node traversed by the setup message. The identifier of the assigned wavelength is recorded in the setup message. If no channel is available on any link, the setup fails, and a notification message is returned to the first hop node.

A new explicit lightpath could be also established through a lightpath setup message which contains the specified route. This route would be determined by the first hop node that could communicate with a higher level network management function. That is, the egress O-LSR has the information of the full network topology and the available resources on every link. These are obtained and updated via IGP (Interior Gateway Routing Protocol) LSAs (Link State Advertisements). IGP extensions (*e.g.*, OSPF and IS-IS) [11], [12] should carry information about the physical state of the fibers in the network. It is necessary to add optical LSA elements such as bit error rate to the IGP in order to support lightpath routing computation.

The explicit route consisting of nodes would be carried in the payload of messages used by one of the available label distribution protocols [13],[14]. Signaling protocols such as the RSVP-TE and the CR-LDP are being extended with objects that provide sufficient details to establish reconfiguration parameters for OXC switch elements. These protocols can do source routing by consulting a TE database. In the case of each protocol, MPLS must operate in the downstream-on-demand distribution mode with ordered control in order for explicit routing to function properly. In the mode, label requests always proceed from the ingress to the egress sequentially and label mappings originate from the egress and propagate in order toward the ingress. In CR-LDP, the *Label Request* message is used to request a wavelength assignment for a specified FEC and CR-LSP. This message is initially issued at the ingress O-LSR and is propagated to the egress O-LSR. A *Label Mapping* message is then generated by the egress OXC. Upon receiving a Label Mapping message from a downstream OXC, an intermediate upstream OXC connects an input port to the output port indicated by the optical label in the message. In contrast to how CR-LDP operates, RSVP is receiver-oriented, uses soft state (*i.e.* it requires periodic transmission of refresh

messages), and encapsulates messages using either raw IP or UDP (User Datagram Protocol) (which means that messages may be lost or received out of order, although the state refresh feature compensates for this to a degree). To establish a connection, a *Path* message is propagated from the source to the destination. The lightpath is set up when a *Resv* message is propagated back to the source from the destination.

A explicit route must be removed when it is no longer needed. To do this, an explicit release request is sent by the edge O-LSR along the lightpath route. Each OXC in the path processes the release message by releasing the resources allocated to the lightpath, and removing the associated state.

3   System Model and ER Procedure

An LSP set-up is requested for a route server which determines the explicit route for the LSP. In case that the LSPs are being set up manually, the request either arrives directly to the server or arrives first at the ingress OXC and then the OXC queries the route server. For determining the explicit route, the route server needs to know the current topology and the available wavelengths. We assume the virtual topology is either known administratively or that a link state routing protocol is operational and that its link-state database is accessible. The algorithm routes the lightpaths over the topology, and assign wavelengths optimally to the various lightpaths. That is, the algorithm keeps track of available wavelengths and handles properly the limited number of the available wavelengths. This assignment problem has been shown to be NP-hard in [15].

We consider a network consisting of N OXCs. Each node is assumed to have a fixed number of ports. A subset of these nodes are assumed to be ingress/egress O-LSRs between which lightpaths can be set up. We assume that the average traffic demand from one edge OXC to another is known. This demand is measured by ISPs, or in the case of Virtual Private Networks (VPNs), specified by customers as the bandwidth requirement for the virtual connection.

Let $G(\mathcal{N}, \mathcal{L})$ describe the given physical network, where $\mathcal{N}$ is the set of N nodes and $\mathcal{L}$ the set of links (*i.e.* fibers) connecting the nodes. Let $\mathcal{K}$ be the set of traffic demands belonging to the DS service class between a pair of edge O-LSRs. Each request $k \in \mathcal{K}$ is defined by the ordered triple $(I_k, E_k, \varepsilon_k)$, where I_k is the ingress OXC, E_k is the egress OXC, and ε_k is the delay limit required for request k. A new LSP can be routed along a given link only if the delay QoS requirement is satisfied.

We now define some of the notations and parameters used.

- C: The capacity of each wavelength on a fiber (bits/sec or pkts/sec).

- $\{D_{ij}\}$: The propagation delays from node i to node j, $i \neq j$. which is proportional to the fiber distance between the two nodes.

- P_n: If a node n $(n = 1, \ldots, N)$ has P_n of ports, clearly, at most $\sum_n P_n$ wavelengths are needed to realize any possible virtual topology. The exact number of wavelengths that are needed to implement a particular virtual topology depends on the physical topology of the network and can be much smaller than $\sum_n P_n$ through wavelength reuse. The number of transmitters and receivers at node $n(n = 1, \ldots, N)$, are defined as $P_n^{(t)}$ and $P_n^{(r)}$, respectively.

- W_{ij}: The number of wavelengths per link in the virtual topology between the nodes i and j for all i and j,

6

- λ_{ij}^{IE}: The traffic from ingress I to egress E that flows over an intermediate virtual link between node i and node j.

- $\lambda^{I_k E_k}$: The average flow associated with the kth traffic demand from the DS service class requesting an ER-LSP set-up.

- $\{x_{ij}^{(v)}\}$: A set of logical variables that indicate how the nodes in the network are linked in the virtual topology:

$$x_{ij}^{(v)} = \begin{cases} 1 & \text{if the virtual topology has a direct fiber link from node } i \text{ to node } j, \\ 0 & \text{otherwise} \end{cases}$$

where $i, j = \{1, 2, \ldots, N\}$ and $i \neq j$.

- $x_{ij}^{(e)}$: The variable related to the ER among the virtual links:

$$x_{ij}^{(e)} = \begin{cases} 1 & \text{if the ER has a lightpath from node } i \text{ to node } j, \\ 0 & \text{otherwise} \end{cases}$$

where $i, j = \{1, 2, \ldots, N\}$ and $i \neq j$.

Further it is assumed that both packet lengths and packet interarrival times at ingress OXC are exponentially distributed.

The constraint conditions are defined as follows. The number of lightpaths originating from and terminating at a node is not more than the node's out-degree and in-degree, respectively. Thus, only one lightpath per port can be setup at each node:

$$\sum_j x_{ij}^{(v)} \le P_i^{(t)}, \qquad \sum_i x_{ij}^{(v)} \le P_j^{(r)} \qquad \text{for all } i, j. \tag{1}$$

(In many cases, we will have $P_i^{(t)} = P_i^{(r)}$.)

There are some constraints related to the traffic flow on the virtual topology for all i and j. First, because we are setting up an ER-LSP, the traffic demand $\lambda^{I_k E_k}$ is not bifurcated at any point in the network. Thus we can write the total flow on the simplex link from node i to node j as the superposition of the existing traffic and the new flow associated with the ER-LSP:

$$\lambda_{ij} = (x_{ij}^{(v)} \sum_{I,E} \lambda_{ij}^{IE}) + x_{ij}^{(e)} \lambda^{I_k E_k} \qquad \text{for all } i, j. \tag{2}$$

Second, the traffic flowing into an OXC should be equal to that flowing out of the OXC for any OXC other than the ingress and egress OXCs for each flow k:

$$\sum_j x_{ij}^{(e)} - \sum_j x_{ji}^{(e)} = \begin{cases} 1, & i = I_k \\ -1, & i = E_k \\ 0, & \text{else} \end{cases} \tag{3}$$

In the above equation, it is also meant that the ER-LSP should be the indicated edge switches I_k and E_k.

Third, traffic flowing through a link can not exceed the total link capacity:

$$\lambda_{ij} \leq W_{ij}C. \tag{4}$$

For the constraint Eq. 4, the layer 3 port throughput can be considered. When the traffic flowing through a link is going forward to an intermediate O-LSR, the traffic demand should not be larger than the sum of the maximum throughput supported by IP router port:

$$\lambda_{ij} \leq W_{ij}C\{1 + (\alpha - 1)Q_j\}, \tag{5}$$

where $\alpha \leq 1$ denotes the maximum layer 3 port throughput rate and Q_j is used to indicate the layer 3 routing capability of the node as follows:

$$Q_j = \begin{cases} 0, & \text{Node } j \text{ has no layer 3 processing} \\ 1, & \text{Node } j \text{ has layer 3 processing} \end{cases} \tag{6}$$

If the link between i and j is not part of the ER-LSP, no traffic associated with the new flow can exist on that link. This constraint can be expressed as

$$\lambda_{ij}^{I_k E_k} = x_{ij}^{(e)} \lambda^{I_k E_k}. \tag{7}$$

In addition, the ER-LSP cannot be set up between two nodes if there is no virtual link connecting them:

$$x_{ij}^{(e)} \leq x_{ij}^{(v)}. \tag{8}$$

A packet traversing the explicit route experiences an end-to-end delay

$$d = d_p + d_n, \tag{9}$$

where $d_p = \sum_{i,j} x_{ij}^{(e)} \lambda_{ij}^{I_k E_k} D_{ij}$ denotes the propagation delay through the LSP and d_n the waiting time in the node. In optical network, the packet transmission delay and queueing delay in an optical buffer might be neglected. So, for d_n, the waiting and processing time should be considered in the intermediate O-LSR which could be the bottleneck on the LSP, since the packets need to be processed in layer 3.

Since there can be an arbitrary number of intermediate O-LSRs on the LSP, we define the vector λ_m^{IE} as aggregate input rate to the mth ($m = \{1, 2, \ldots, M\}$) intermediate O-LSR. So, $M/M/1$ queueing results can be applied to each intermediate OXC which performs layer 3 function, by employing the independence assumption on interarrivals as in Fig. 2. That is, $\lambda_m^{IE}(t)$ is aggregate arrival rate measured in the mth O-LSR at an instant t. The variable μ_m^{IE} denotes the service rate in each intermediate node and the variable $\rho_m^{IE} = \lambda_m^{IE}/\mu_m^{IE}$ gives the traffic intensity.

Then, for this k^{th} traffic flow belonging to the DS class, the objective function for delay can be designed as

$$f_d(x_{ij}^{(e)}) = (\sum_{i,j} x_{ij}^{(e)} \lambda^{I_k E_k} D_{ij}) + \tau, \tag{10}$$

where $\tau = \sum_{m=1}^{M} \frac{\rho_m^{IE}}{(1 - \rho_m^{IE})\lambda_m^{IE}}$ is the average layer 3 processing delay seen by the traffic. We can rewrite using the local connectivity and processing rate variables as

$$\tau = \sum_{i,j} \frac{x_{ij}^{(e)} Q_j}{\mu_j - (\sum_\ell \lambda_{\ell j} + \lambda^{I_k E_k})}. \tag{11}$$

In Eq. 10, the second term τ will be zero if there is no layer 3 control. The function given by Eq. 10 should be minimized in order to support the delay QoS requirement by the kth flow from class DS. If the minimized value of Eq. 10 could not satisfy the requested QoS by class DS, the values computed from the minimization would not be applied to the variable $x_{ij}^{(e)}$.

In addition to the above constraints (Eq. 2 - 7), another two constraints which are related to QoS and layer 3 processing rate should be defined. One constraint for QoS can be defined as

$$0 \leq f_d(x_{ij}^{(e)}) \leq \varepsilon_k. \tag{12}$$

Then, the requested QoS by k^{th} flow would be satisfied. The other constraint for the layer 3 processing rate can be expressed as

$$\mu_j \geq x_{ij}^{(v)} \sum_i \lambda_{ij}^{IE} + x_{ij}^{(e)} Q_j \lambda^{I_k E_k} \qquad \text{for all } I, E. \tag{13}$$

According to the values of the $x_{ij}^{(e)}$, the ER-LSP would be set up. The multicommodity network flow problem with integer constraints is generally known to be NP-hard [16]. The k disjoint route problem which is NP-hard in [16] can be dealt with the same as that the k distinct ingress, egress node pairs find k mutually link-disjoint routes.

Whenever a new traffic flow belonging to the DS service class requests an explicit route, the virtual lightpath will be configured. The procedure will be proceeded as followings:

Step 0: Check if there are any ER-LSP set-up requests of kth traffic flow of DS service class in the set $\mathcal{K}$;
Step 1: if a transmitter or a receiver is not available at I_k or E_k, respectively
      then go to Step 3;
Step 2: if there is any ER-LSP to minimize Eq. 10
      then set up the lightpath between I_k and E_k;
      else renegotiate on this flow or go to Step 3;
Step 3: Block this ER-LSP request;

As the network loading varies over time, the consideration of the optimal route selection would likely result in the reconfiguration of lightpath routes being required. Although frequent lightpath reroutings may not be acceptable, a limited number of lightpath reroutings could improve the network performance, supporting the requested QoS of future traffic while maintaining the QoS of the traffic that the network is already supporting. For restoration, rerouting would also have to be performed within the time limits set for restoration, which may impose tight constraints on the amount of time allowed to establish restoration paths. To satisfy the requirements of these diverse routing, rerouting and restoration as well as traffic engineering, explicit routing is necessary for constructing lightpaths. The route on which a new lightpath is to be established is specified by an information (Object/TLV) contained in the lightpath setup message. This route is typically be chosen by the ingress O-LSR, but it could be determined by a higher level network management system. The route may be specified either as a series of routers/OXCs, or in terms of the specific links used. Therefore, the above mechanism performs the calculation of primary and restoration lightpath routes on-line as the individual requests arrive. These lightpaths could be computed all at once by doing an offline calculation that accounts for all the pending requests. And also, In the initial configuration stage where there is no configured virtual topology, the appropriate virtual lightpath could be found by repeating the above procedure. This procedure is applied to the traffic with the highest delay

limit among all initial traffic demands of the DS service class at all ingress O-LSRs being done so to the traffic with the next delay limit in turn. Like setting up an ER-LSP, the virtual topology would be configured for the traffic demand of the DS service class at each ingress OXC by minimizing the objective function of delay. By finding the values of all the elements of the set $\{\lambda_{ij}^{IE}\}$, we obtain a full set of routing assignments for all the traffic in the optical network.

For BE service class, the objective function to be minimized is defined as

$$f_c(x_{ij}^{(v)}) = \max_{i,j} \lambda_{ij}. \tag{14}$$

This objective function is derived by finding the link that suffers from maximum congestion. Therefore, for the BE service class, the proposed mechanism focuses on distributing the loading within the entire network rather than on meeting a desired QoS.

4   Performance Evaluation from Simulation

The performance of the proposed algorithm and that of the SP (Shortest Path) algorithm have been analyzed using simulations. In this section, we describe a simulation setup used to validate our algorithm in providing service requested by the DS class and then compare the performance of the two algorithms. As simulation tool, we used the simulator MERLiN [17] which was developed for WDM network simulations, for the analysis with our proposed algorithm described earlier. The simulation tests were carried out on a model of the network shown in Fig. 3 [18] with a maximum of 4 wavelengths are available for use on each link. For this topology model, two cases were tested. While in one case (we call it TEST 1), the values of the Q_j are assumed to be 1 for all the edge nodes, in the other case, the values are 1 only for j=5,8,9,12 and 14 (TEST 2). As for the traffic generation, we assume that every traffic flow requesting ER-LSP is generated according to a Poisson process with two values, 200 and 1000. Each simulation run continues till 5000 LSP setup requests are generated. We assume that the time between lightpath set up and teardown is exponentially distributed with a mean of 1. We generated the lightpath setup requests and the traffic flows at all edge nodes. When an ER-LSP setup request arrives at a source node, the destination is chosen randomly among all the edge nodes except the ingress with the setup request. 10% of the total traffic flows belongs to the DS service class. Table 1 shows an numerical example of the network parameters used in simulation.

Fig. 4 and 5 show the blocking performance for overall traffic and for only the DS traffic, respectively, with LSP setup request rate, $1/\lambda = 0.001$ (The unit over x axis is global time unit for event management in the simulator). We observe that over overall routes, improvements in blocking probability can be achieved during overall simulation duration. Especially, for the DS traffic, the blocking performance is significantly improved by selecting the path with the minimum delay, rather than the shortest path like in Fig. 12. Fig. 6 and 7 also show a reduction in the blocking probability when $1/\lambda = 0.005$. This shows that our algorithm utilizes the unused wavelength.

Similar results are observed for the other set of simulations where the location of the O-LSRs is changed, i.e. the layer 3 processing is removed at the nodes 5, 8, 9, 12 and 14. The blocking probability for TEST2 is illustrated in Fig. 8, 9, 10, and 11. As can be seen from the graphs, our proposed algorithm performed better with respect to the blocking performance, for both $1/\lambda = 0.001$ and 0.005. Compared to the results in Fig. 8 and 9, only the scale of the blocking probability has been reduced in Fig. 10 and 11 since the applications request explicit route

less frequently at edge O-LSR. And as can be seen in the graphs for only the DS traffic, the blocking probability is higher than that of overall traffic in order to guarantee the delay QoS requested by the DS service class.

The improvements in blocking performance means that the proposed algorithm utilizes network resources more efficiently. That is, for the DS traffic, an attempt is made firstly to establish a new connection using the path with the minimum delay. If the connection cannot be established using that path, then the proposed algorithm chooses the shortest path as the SP algorithm does on condition that the requested delay QoS is satisfied.

The average delay associated with both algorithms is presented in Table 2, measured during the overall simulation time for the two tests (TEST 1 and TEST 2) where one performs layer 3 processing at all the edge nodes and the other one does so at only the 5 nodes listed above. Table 2 shows that the delay performance could be improved without performance degradation in blocking rate. This results from considering layer 3 processing delay to set up an LSP as can be seen in Fig. 12. In the Fig. 12, the two algorithms chose different paths between source and destination pairs (2, 9) and (14, 12). In other words, while the proposed algorithm selected the routes, $(2\rightarrow3\rightarrow7\rightarrow9)$ and $(14\rightarrow10\rightarrow11\rightarrow12)$, the SP algorithm chose the routes $(2\rightarrow1\rightarrow4\rightarrow9)$ and $(14\rightarrow13\rightarrow12)$. The SP algorithm takes longer delay than our algorithm because it experiences layer 3 processing at nodes 1 and 4 for the connections between nodes 2 and 9. Note that our algorithm also takes better delay performance for the connections between nodes 14 and 12 even though it contains one more hop over the path. From the Table 2 and the Fig. 12, it can be known that the proposed algorithm improves the delay performance by selecting the lightpath without layer 3 processing. After increasing layer processing rate (1.5 Gbps) at nodes 1, 4, and 13, the simulation tests was performed again to look into a change of delay with $1/\lambda=0.001$. In TEST 1, the average delays for the proposed algorithm and the SP algorithm are 0.249 and 0.290 msec, respectively. Meanwhile, in TEST 2, the average delays are measured as 0.275 and 0.303 msec, respectively for the proposed algorithm and the SP algorithm. While the improvements in TEST 1 and TEST 2 are 14.1% and 9.2%, respectively, in the case that the processing speed is 1.5 Gbps, 45.3% and 38.8% are improved in TEST 1 and TEST 2, respectively, when the speed is 0.5 Gbps. Examining the connections between nodes 2 and 9 for TEST 1, and the connections between nodes 14 and 12 for TEST 2, the same kind of results are obtained as can be seen in Table 2: for TEST 1, the improvement in delay performance increases from 12.7% to 44.8% over the path (2-9), while the delay is improved from 5.4% to 34.6% over the path (14-12) for TEST 2. This is because the slower each O-LSR processes the data at layer 3, the longer it takes the processing delay. It is therefore important to study the effects of layer 3 processing delay on network performance.

5 Conclusion

In this paper, we proposed an optimization algorithm to support the requested delay QoS of LSP requests in an optical network that uses lambda labeling to switch lightpaths. This algorithm uses the current state of the network to determine the delay associated with each possible path, and then chooses the path with the minimum total delay. A novel feature of this approach is that it accounts for the delay encountered by packets that require layer 3 processing at subnetwork edge nodes. Using a model of a meshed optical backbone network, we used the MERLiN tool to demonstrate that traffic assigned to LSPs with our algorithm experiences less delay and lower blocking probability than traffic that is assigned using a standard algorithm such as shortest path first.

References

[1] R. Callon et al., "A Framework for Multiprotocol Label Switching", *Internet Draft*, draft-ietf-mpls-framework-05.txt, Sep. 1999.

[2] E. Rosen et al., 'Multiprotocol Label Switching Architecture", *RFC 3031*, Jan. 2001.

[3] G. Apostolopoulos et al., "QoS Routing Mechanisms and OSPF Extensions," *RFC 2676*, Aug. 1999.

[4] N. Golmie, T. D. Ndousse, and D. Su, "A Differentiated Optical Services Model for WDM Networks," *IEEE Communications Magazine*, vol.38, no.2, pp.68-73, Feb. 2000.

[5] R. Ramaswami and K. N. Sivarajan, *Optical Networks: A Practical Perspective*, Morgan Kaufmann, 1998.

[6] H. Harai, F. Kubota, and H. Nakazato, "Design of Reconfigurable lightpaths in IP over WDM Networks," *IEICE Transactions on Communications*, vol.E83-B, no.10, pp.2234-2244, Oct. 2000.

[7] B. Rajagopalan et al., "IP over Optical Networks: A Framework", *Work in Progress*, draft-many-ip-optical-framework-02.txt, Jul. 2000.

[8] P. Ashwood-Smith et al., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," *Work in Progress*, draft-many-gmpls-architecture-00.txt, Feb. 2001.

[9] D.O. Awduche et al., "Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control With Optical Crossconnects", *Work in Progress*, draft-awduche-mpls-te-optical-03.txt, Jul. 2000.

[10] B. Jamoussi et al., "Constraint-Based LSP Setup using LDP", *Internet Draft*, draft-ietf-mpls-cr-ldp-05.txt, Feb. 2001.

[11] D. Katz, and D. Yeung, "Traffic Engineering Extensions to OSPF," *Work in Progress*, draft-katz-yeung-ospf-traffic-01.txt, Aug. 2000.

[12] H. Smit, and T. Li, "IS-IS Extensions for Traffic Engineering", *Work in Progress*, draft-ietf-isis-traffic-01.txt, Sep. 2000.

[13] P.A. Smith, et al., "Generalized MPLS Signaling - CR-LDP Extensions," *Work on Progress*, draft-ietf-mpls-generalized-cr-ldp-01.txt, Mar. 2001.

[14] P.A. Smith et al., "Generalized MPLS Signaling - RSVP-TE Extensions," *Work in Progress*, draft-ietf-mpls-generalized-rsvp-te-01.txt, Mar. 2001.

[15] I. Chlamtac, A. Ganz, and G. Karmi, "Lightnets: Topologies for High Speed Optical Networks," *IEEE/OSA Journal of Lightwave Technology*, vol.11, no.5/6, pp. 951-961, May/Jun1993.

[16] M. Garey, and D. Johnson, *Computers and Intractability*, W.H. Freeman and Company, 1979.

[17] NIST, http://w3.antd.nist.gov/Hsntg/prd_merlin.html.

[18] K. Kar et al., "Minimum Interference Routing of Bandwidth Guaranteed Tunnels with MPLS traffic Engineering Applications," *IEEE JSAC Commun.* vol.18, no.12, Dec. 2000.

Table 1: Parameter Values

Parameter Symbol	Value
W_{ij}	4
$P_i^{(t)}, P_i^{(r)}$	4
C	10 Gbps
D_{ij}	0.05 msec
$\lambda_{IE,k}$	2.5 Gbps
ε	0.5 msec

Table 2: Delay Performance(Unit: msec)

TEST 1

$1/\lambda$	Proposed algorithm	SP algorithm
0.001	0.292	0.477
0.005	0.266	0.424
Path (2-9): Processing speed=0.5 Gbps		
0.001	0.275	0.498
Path (2-9): Processing speed=1.5 Gbps		
0.001	0.254	0.291

TEST 2

$1/\lambda$	Proposed algorithm	SP algorithm
0.001	0.228	0.417
0.005	0.215	0.371
Path (14-12): Processing speed=0.5 Gbps		
0.001	0.315	0.482
Path (14-12): Processing speed=1.5 Gbps		
0.001	0.294	0.311
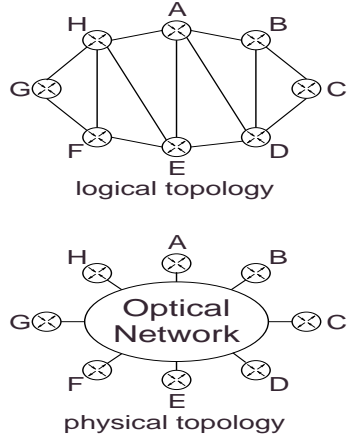
Figure 1: Sample optical network with IP routers located at the edge. Lightpaths can be created across the optical network to form the virtual topology shown at the top of the figure.
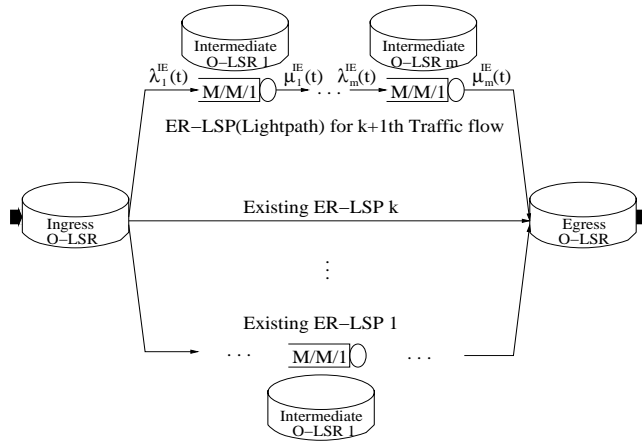
Figure 2: Delay when intermediate nodes exist
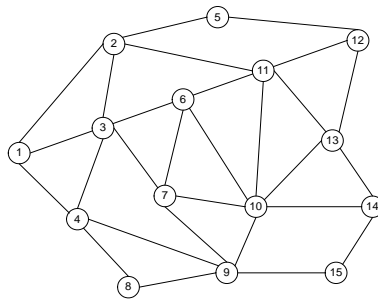
Figure 3: Network topology
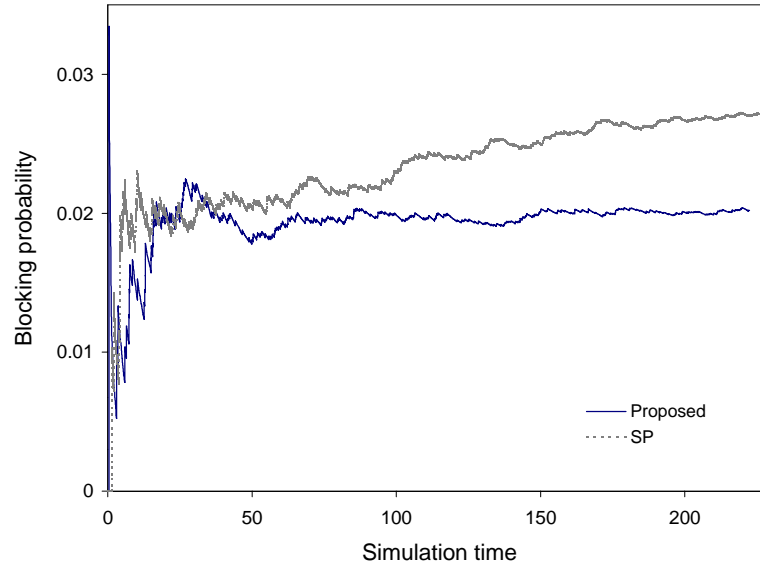
Figure 4: Blocking probability(TEST 1; Overall traffic; $1/\lambda$=0.001)
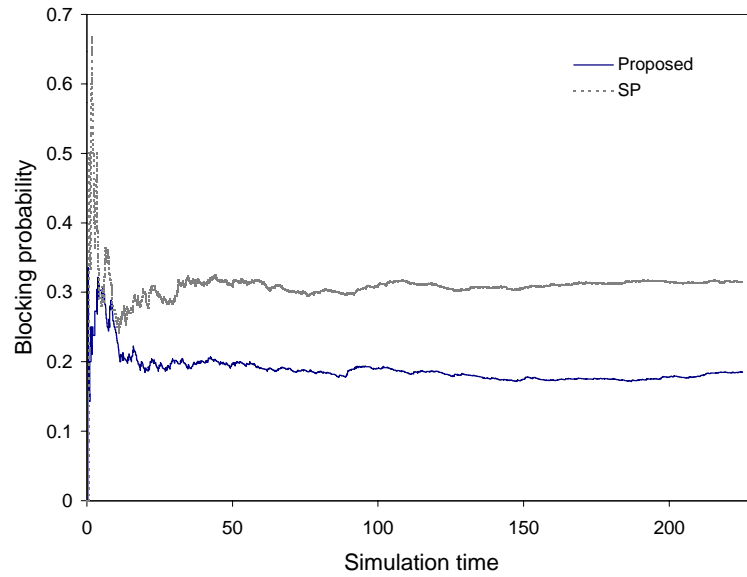
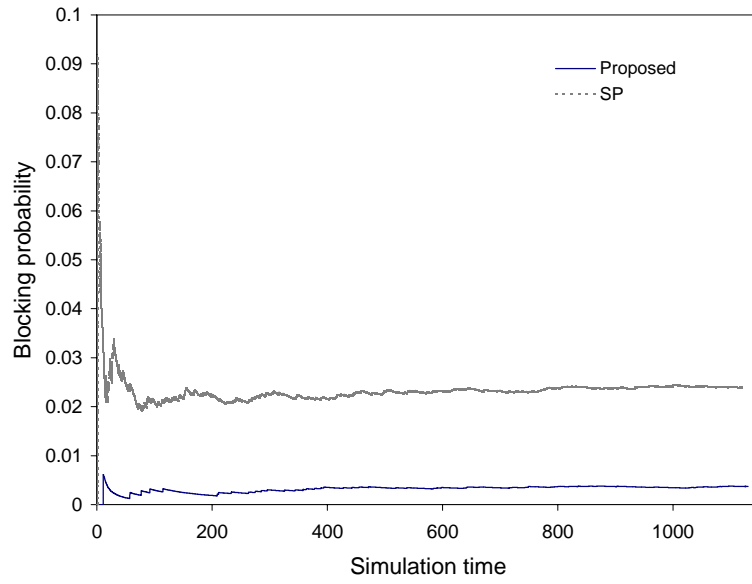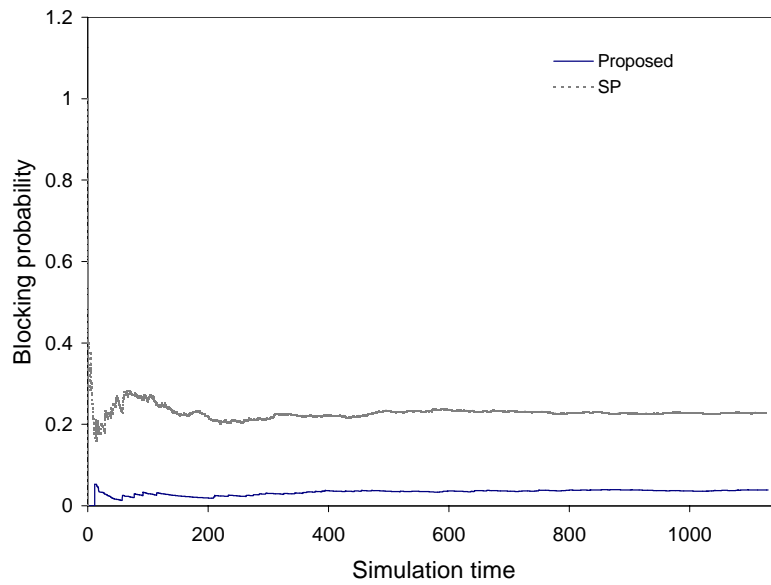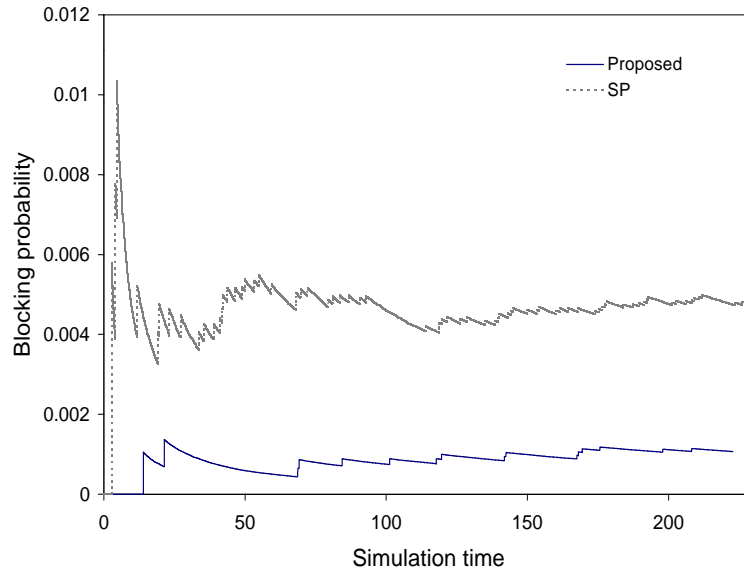Figure 5: Blocking probability(TEST 1; DS traffic; $1/\lambda$=0.001)

Figure 6: Blocking probability(TEST 1; Overall traffic; $1/\lambda$=0.005)

Figure 7: Blocking probability(TEST 1; DS traffic; $1/\lambda$=0.005)

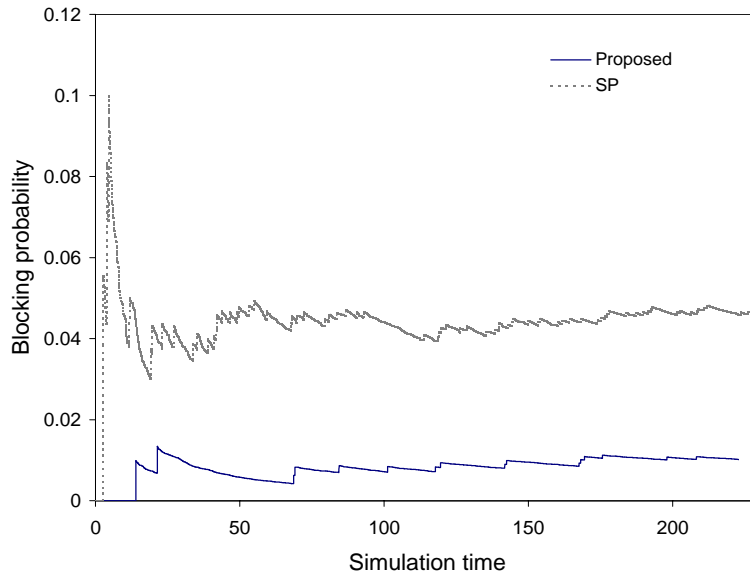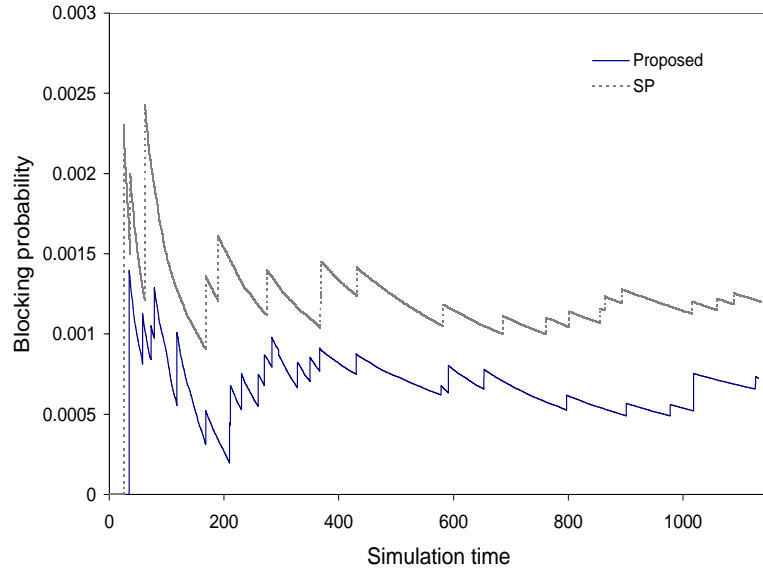Figure 8: Blocking probability(TEST 2; Overall traffic; $1/\lambda$=0.001)

Figure 9: Blocking probability(TEST 2; DS traffic; $1/\lambda$=0.001)

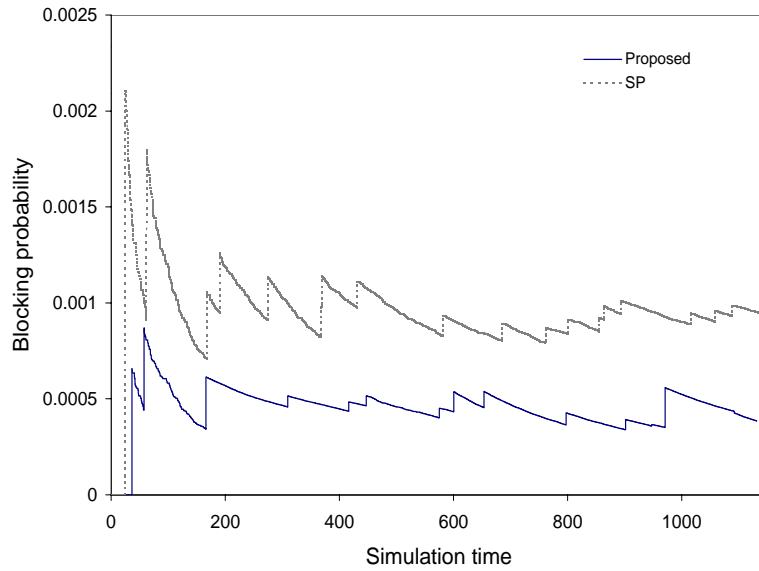Figure 10: Blocking probability(TEST 2; Overall traffic, $1/\lambda$=0.005)

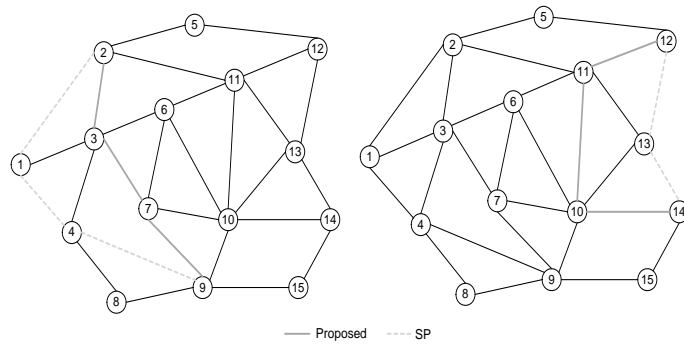Figure 11: Blocking probability(TEST 2; DS traffic, $1/\lambda$=0.005)

18

Figure 12: Path setup